SUPPLEMENTARY MATERIALS

Figure 1

Schematic diagram of the organization of domains and modules in NRPS and modular PKS. The minimal core module in case of NRPS consists of adenylation (A) domain, condensation (C) domain and thiolation or peptidyl carrier protein (T or PCP) domain. The optional modifying domains in NRPS modules are epimerization (E) domains for stereoisomerization of monomers and methyltransferase (M) domains responsible for N-methylation of the peptide backbone. An acyltransferase (AT) domain, an acyl carrier protein (ACP) and a ketoacyl synthase (KS) domain constitute the core domains of PKS modules. In case of PKS, the optional modifying domains are ketoreductase (KR), methyl transferase (O-MT), dehydratase (DH) and enoyl reductase (ER).

Figure 2

Screen shot from CHSDB depicting the chemical structure of the starter, extender and final product for acridone synthase. It also shows interfaces for extracting FASTA sequence, identifying putative active site residues and comparing the 32 active site residues of acridone synthase with those of other CHS proteins.

Figure 1.

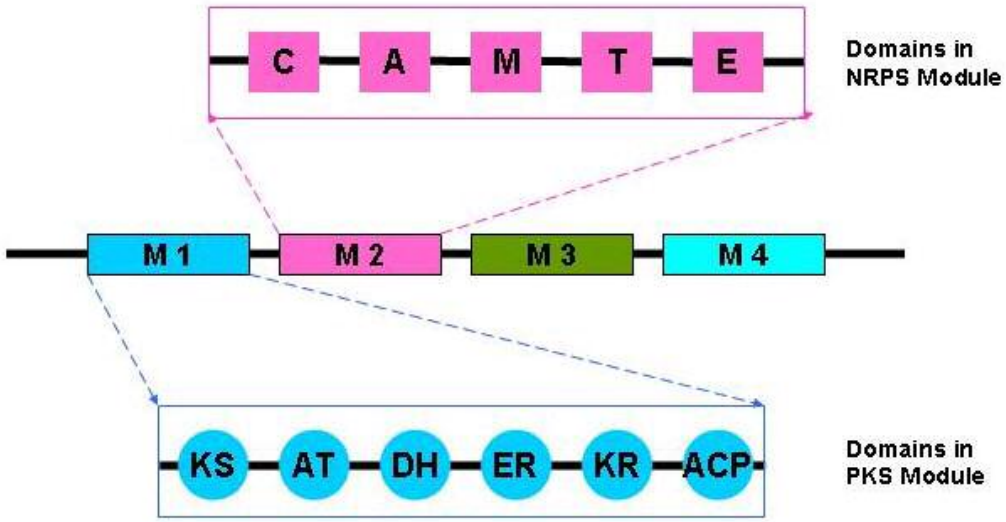
C
A
M
T
E
Domains in
NRPS Module
M 1
M 2
M 3
M 4
KS
AT
DH
ER
KR
ACP
Domains in
PKS Module

Figure 2.

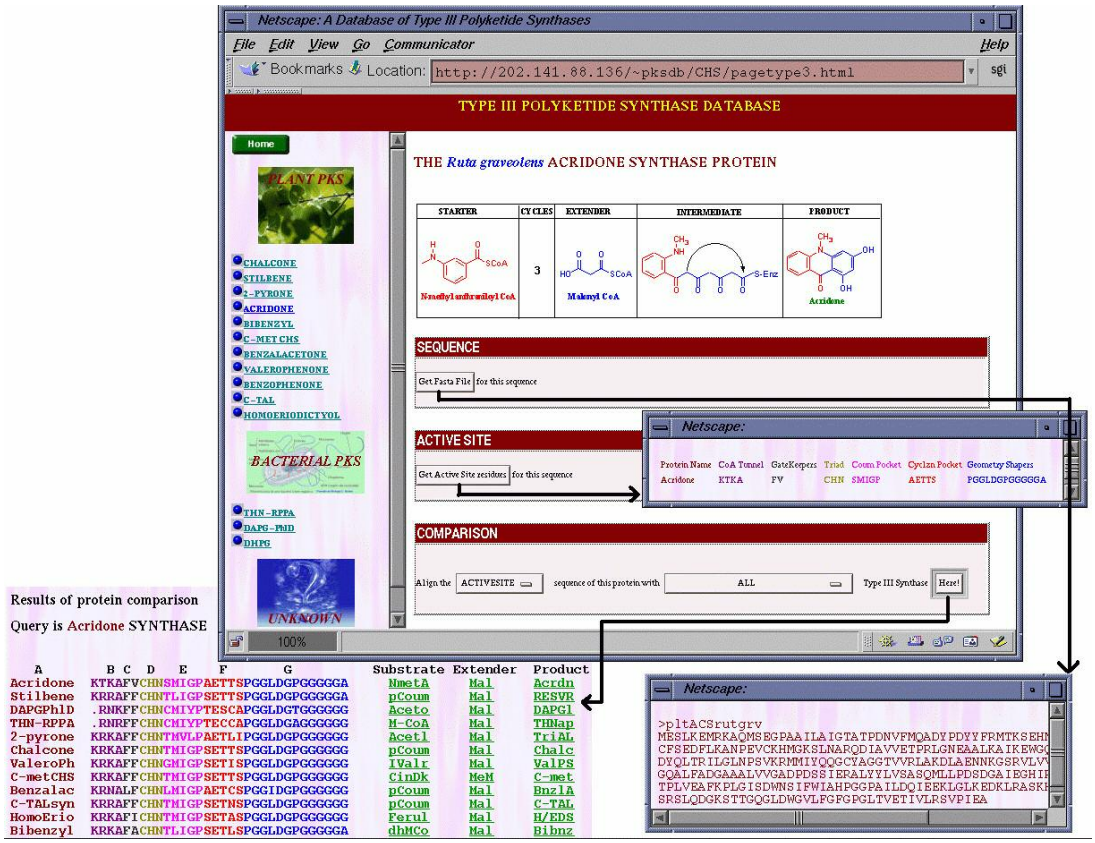
Netscape: A Database of Type III Polyketide Synthases
File Edit View Go Communicator Help
Bookmarks Location: http://202.141.88.136/~pksdb/CHS/pagetype3.html
TYPE III POLYKETIDE SYNTHASE DATABASE
Home
PLANT PKS
CHALCONE
STILBENE
2-PYRONE
ACRIDONE
BIBENZYL
C-MET CHS
BENZALACETONE
VALEROPHENONE
BENZOPHENONE
C-TAL
HOMOERIODICTYOL
BACTERIAL PKS
THN-RPPA
DAPG-PhID
DHPG
UNKNOWN
THE Ruta graveolens ACRIDONE SYNTHASE PROTEIN
STARTER
CYCLES
EXTENDER
INTERMEDIATE
PRODUCT
3
Malonyl CoA
Acridone
SEQUENCE
Get Fasta File for this sequence
ACTIVE SITE
Get Active Site residues for this sequence
Netscape:
Protein Name CoA Tunnel GateKeepers Triad Coum Pocket Cyclzn Pocket Geometry Shapers
Acridone KTKA FV CHN SMIGP AETTS PGGLDGPGGGGGA
COMPARISON
Align the ACTIVESITE sequence of this protein with ALL Type III Synthase Here!
Netscape:
>pltACSrutgrv
Results of protein comparison
Query is Acridone SYNTHASE
A B C D E F G Substrate Extender Product
Acridone KTKAFVCHNSMIGPAETTSPGGLDGPGGGGGA NmetA Mal Acrdn
Stilbene KRRAFFCHNTLIGPSETTSPGGLDGPGGGGGA pCoum Mal RESVR
DAPGPhlD .RNKFFCHNCMIYPTESCAPGGLDGTGGGGGG Aceto Mal DAPGl
THN-RPPA .RNRFFCHNCMIYPTECCAPGGLDCAGGGGGG M-CoA Mal THNap
2-pyrone KRKAFFCHNTMVLPAETLIPGGLDGPGGGGGG Acetl Mal TriAL
Chalcone KRKAFFCHNTMIGPSETTSPGGLDGPGGGGGG pCoum Mal Chalc
ValeroPh KRKAFFCHNGMIGPSETISPGGLDGPGGGGGG IValr Mal ValPS
C-metCHS KRKAFFCHNTMIGPSETTSPGGLDGPGGGGGG CinDk MeM C-met
Benzalac KRNALFCHNLMIGPAETCSPGGIDGPGGGGGG pCoum Mal BnzlA
C-TALsyn KRRAFFCHNTMIGPSETNSPGGLDGPGGGGGG pCoum Mal C-TAL
HomoErio KRKAFICHNTMIGPSETASPGGLDGPGGGGGG Ferul Mal H/EDS
Bibenzyl KRKAFACHNTLIGPSETLSPGGLDGPGGGGGA dhMCo Mal Bibnz

Table 1.
Summary of domain identification in the test set. Results for NRPS domains are given in Table 1a, while 1b lists the results for PKS domains. For each domain type, the table lists the total number of such domains present in all the clusters included in the test set, number of domains detected by NRPS-PKS and the list of domains missed or over predicted by NRPS-PKS. For A and AT domains, table also gives the number of such domains for which correct specificity is predicted by NRPS-PKS. The names of the clusters included in the test set and results given by NRPS-PKS for each one of them are available at http://www.nii.res.in/nrps-pks/benchmark.html .

a

Domain	Literature	SEARCHNRPS	Missed	Over Pred	Specificity
C	108	106	**2**: ennia mod 1, vibri mod 2	0	
A	90	90	0	0	Exact 52 Similar 16
T	110	105	**5**: ClmH0 mod 1, myxot mod 1, 4, 9 peder mod 11	0	
M	6	4	**2**: anaba mod 5 barba mod 6	0	
TE	15	15	0	0	

b

Domain	Total	SEARCHPKS	Missed	Over Pred	Specificity
KS	70	69	**1**: leina mod 4	0	
AT	56	56	0	0	52
DH	38	36	**4**: leina mod 4 mupir mod 1, 2, 3	**2**: leina mod 6 ansam mod 7	
ER	17	17	0	0	
KR	59	61	0	**2**: ansam LD gelda LD	
ACP	79	85	0	**6**: ansam mod 5,7 leina mod 8 mupir mod 2, 5	
TE	4	3	**1**: monensin TE II	0	